Study on Selecting the Surgical methods for Patients with Knee Osteoarthritis based on BP Neural Network

Zhang Haohua¹, Wang Jiaying², Zhang Kuan²,³, Yan Songhua²,³*

Abstract
Knee Osteoarthritis (KOA) seriously affects life quality of patients. Knee arthroplasty, including total knee arthroplasty (TKA) and Unicompartmental Knee Arthroplasty (UKA), is the main and effective treatment method for KOA patients. When selecting the surgical method, the physicians often make decision based on experiences and patient feeling by conducting an interview, KSS score and X-ray test. The diagnosis process is lengthy and subjective. Thus, the aim of this study is to establish a prediction model and explore the methods of selecting surgical way for KOA patients. One hundred and eighty-three patients with moderate and severe KOA were chosen as our subjects. BMI, body weight, age, gender, KSS score and the used surgical way were recorded. X-ray parameters were obtained. Range of Motion (ROM) of knee joint was measured by a wearable motion monitoring system. Back propagation neural network was applied to get the weight of various parameters affecting the selection of surgical methods and a prediction model was established. The results showed that the lateral space of knee joint had the largest relative weight. BMI, age, gender, medial space of knee and ROM had moderate weight; osteophyte or not, weight, and pain scores had less weight. The total accuracy of the established surgical prediction model reached 92.97%. So, when surgeon judges the surgical method, the lateral space of knee is an important parameter. At the same time, the surgeon may combine age, body weight, sex, ROM and medial space. BMI, pain score, and osteophytes or not can just be used as the reference.

Keywords: Knee Osteoarthritis (KOA); Surgical way; Machine learning; Back Neural Network; Wearable device

Introduction
Knee osteoarthritis (KOA) is a common progressive multifactorial joint disease and is characterized by chronic pain and functional disability [1] because of sclerosis of articular cartilage or osteoproliferation, narrowing of the knee joint space [2]. KOA patients have a greater impact on quality of life for the limited joint mobility. Moreover, KOA is more common in middle-aged and elderly patients with poor recovery ability, so it is necessary to cooperate with drug therapy or surgical treatment to prevent further wear of articular cartilage and relieve the pain for KOA patients [3-5].

Knee arthroplasty is the main and effective treatment method for patients with mid- and end-stage KOA in clinical surgical treatment [6], which aims to restore the function of knee joint and improve the quality of life of patients [7,8]. Total Knee Arthroplasty (TKA) and Unicompartmental Knee Arthroplasty (UKA) are two kinds of operation method for treating...
moderate and severe KOA [9]. UKA was reputed to show lower morbidity and mortality than TKA [10]. Compared with TKA, UKA showed a lot of advantages such as smaller incision, less soft tissue injury and blood loss, preservation of bone stock, normal knee kinematics, lower infection rate, shorter hospital stay, faster recovery, and greater range of motion [11,12]. However, there are more contraindications for UKA. For patients with severe joint injury, TKA may be more suitable for them.

It is expected that the elderly population in China will exceed 300 million in 2025, and it will continue to rise [13]. KOA has become a common disease in the elderly, the number of patients is also increasing. However, when doctors make diagnosis, they need to interview the patient based on the result of X-ray and KSS score and decide to use UKA or TKA according to experience and patient opinion. The diagnosis process is time-consuming, and contains certain subjective factors. There is no objective method to solve this problem at present. Artificial intelligence (AI) is characterized by great stability, efficiency and objective accuracy. Proper use of AI to establish a prediction model can provide the weight of parameters for doctors to refer to, and give relatively objective prediction opinions, which will greatly improve the work efficiency of doctors.

Machine learning is the most common technology used in current AI system. It can use algorithms to explain and analyze data, learn from it, and then make decisions and predictions from real-world events. In 1940s, Warren McCulloch and Walter Pitts [14] proposed the first mathematical model of artificial neuron (MP neuron model). Researchers have listed several machine learning algorithms for predicting and classifying Parkinson's disease, which includes Artificial Neural Network (ANN), K-Nearest Neighbor algorithm (K-NN), Support Vector Machine (SVM), Random Forest Algorithm, Deep Neural Network (DNN), Mtilayer Perceptron, K-means clustering algorithm and Genetic Algorithm (GA). They concluded that SVR was a valuable algorithm for predicting Parkinson's disease, especially for data on gait characteristics [15]. AI and machine learning were also being applied to the research on joint replacement of lower limb [16]. Some researchers have applied adaptive prediction system to gait analysis using neural network algorithm [17]. In the field of orthopedics, there were also studies using AI to diagnose distal radius fractures, hip fractures and osteoporosis [18,19]. However, most of the current studies on knee arthroplasty were focused on analyzing the factors influencing patients' postoperative recovery, without dealing with the factor analysis and evaluation of the surgical choice during preoperative diagnosis.

At present, the most commonly used neural network is linear neural network, BP neural network [20] and convolutional neural network [21]. Linear neural network can solve the most basic linear separable problem, but cannot solve the linear non-separable problem. BP neural network contains multiple hidden layers and can solve linear non-separable problems. So BP neural network has widely used in pulse recognition and diagnosis in medical area [22]. Convolutional neural network is mostly used in image recognition, automatic driving, face recognition and other fields with more complex functions [23-26].

Therefore, the aim of this study was to explore a new method for doctors to select surgical methods for patients with moderate or severe knee osteoarthritis. We applied BP neural network to analyze the indicators and surgical methods used by doctors in clinical diagnosis. Firstly, the weights of various parameters affecting the selection of surgical methods were determined and analyzed by using the basic data, KSS score and the morphology index of the X-ray image. Then a prediction model for selecting surgical methods was established, finally the prediction accuracy by the established model was judged based on the doctor's decision. This method was expected to play an auxiliary role in the clinical diagnosis of doctors, and a more scientific selection of surgical methods was suitable for patients.

**Subjects and Methods**

**Subjects**

A total of 183 KOA patients from the clinic in orthopaedic department of Beijing Jishuitan Hospital were chosen as the subjects in this study. They were going to be performed TKA or UKA for treatment. Patient-related data was collected to establish patient data sets. The patient data set was randomly divided into two groups. Among 183 groups of patient data, 138 groups were used as training data for machine learning neural network training, and the other 45 groups were used as test groups to verify the accuracy of the prediction model. The surgical methods of the patients were discussed by three deputy chief physicians or chief physicians of Beijing Jishuitan Hospital.

Inclusion criteria: (1) all patients were clinically diagnosed as KOA in Beijing Jishuitan hospital, which met the diagnostic criteria of Chinese medical association [27]; (2) patients who underwent knee replacement for the first time on the affected leg; (3) all patients did not suffer from serious organ diseases, such as cardiac dysfunction and renal failure.

Exclusion criteria: (1) patients who did not meet the diagnostic criteria of knee osteoarthritis by Chinese medical association, such as rheumatoid arthritis [28] and traumatic arthritis; (2) patients undergoing lateral knee arthroplasty; (3) patients who did not accept the survey, or those who did not have complete information; (4) patients with mental diseases, consciousness disorders [29].

All subjects participating in this study signed informed consent. The study was approved by the Ethics Committee of Jishuitan Hospital.

**Experiment device**

In this study, a wearable motion monitoring system (iMotion wearable motion monitoring tester) was used for measuring the Range of Motion (ROM). This system consists of data acquisition modules, patient APP, cloud server, doctor APP. The data acquisition module collected kinematics data and transmitted it to the patient APP through Bluetooth. Data acquisition module is a sensor, including motion sensor, microcontroller, power supply and Bluetooth module. The motion sensor consists of accelerometer, gyroscope, and geomagnetic sensor, all of which have three axes. The three-dimensional angle was calculated according to the theory of posture and heading reference system by nine-axis data. Data acquisition modules were fixed on the thigh and tibial anterior region respectively to measure the three-dimensional angle of the knee joint, where the longitudinal angle was ROM. When the patient's knee joint was used before surgery, the patient's ROM can be detected in the doctor app and automatically recorded. The system has proven to be reliable [30]. In this study, this system was used to collect ROM from patients. The system also recorded the patient's height, weight, age, sex, and pain scores, which can be digitized for subsequent studies.

**Parameters**

Parameters include three categories: basic information parameters, KSS score and X-ray image parameters.

Basic information parameters include patient gender (male/female), age (years), weight (kg) and Body Mass Index (BMI, kg/m²).

KSS score includes two items according to the Knee Society Score (KSS) of American Knee Society [31,32], that is, ROM [33] and preoperative pain scores. Knee motion was measured with the iMotion wearable motion monitoring tester. Preoperative pain score was scored according to the complaint pain scale in the KSS score, as shown in Table 1. Total scores of Pain scale are 50 points. Patients were asked to walk on the ground and climb stairs respectively, and the experts scored for them by listening to patients’ complaint. Preoperative pain scores were obtained by adding all scores.

X-ray images of the patient's knee joint on the surgical side were taken and read by doctor. X-ray image parameters included whether there were osteophytes, sclerosis and a narrowing of the lateral compartment and the medial compartment or not.

**Protocol**

**Study procedure**

Firstly, the patient basic data, KSS score and X-ray image data were sorted out and randomly grouped. Then, Age, gender, weight, BMI, pain score, ROM of knee joint, whether osteophyte or not, whether sclerosis or not, whether narrow lateral compartment or not, whether narrow medial compartment or not on X-ray images were input into the neural network, and neural network were trained. Finally, the weight of each parameter was obtained and the prediction model was established.

**Neural Network**

**Datasets pre-processing**

The textual data were replaced with binary data that can be read by the computer. Gender was coded as "1" for "male" and "0" for "female", and "yes" was coded as "1" for osteophyte, sclerotic, narrow medial and lateral compartment, and "0" for "no". The patient was coded as "1" for "TKA" and "0" for "UKA".

The 10 data were imported into a csv file named as "input.csv", and each patient data was coded in the same column for each type of input parameter. The patient's final procedure data was imported into another csv file "named as output.csv" in the same order as the "input.csv" file.

**Datasets creation**

"Input.csv" with 183 patient data was read and a dataset containing 183 10-column and 1-row vectors were created. At the same time, the patient surgery used was read to build a label for each dataset. Each vector was randomly disordered and divided into training and test sets, one hundred and eighty-three for the training set and 45 for the test set.

**Establishment of prediction model**

In this study, ten parameters were included. It is not possible to use linear neural network because of the complex problem. And the convolutional neural network is widely used in the field of image recognition, which has a large amount of calculation and long data processing time. For a medical problem with ten parameters, there is no need to perform such a large amount of computation. Therefore, BP neural network was chosen to be constructed in this study.

The neural network was built by using Keras in Tensorflow AI learning system. BP neural network with input layer (10 neural nodes), 1 hidden layer (32 neural nodes), and output layer (2 neural nodes) was constructed (Figure 1).

**Table 1: KSS pain scale.**

<table>
<thead>
<tr>
<th>Walking</th>
<th>No pain (35 Point)</th>
<th>Mild or occasional pain (30 Point)</th>
<th>Moderate pain (15 Point)</th>
<th>Severe pain (0 Point)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Climbing the stairs</td>
<td>No pain (15 Point)</td>
<td>Mild or occasional pain (10 Point)</td>
<td>Moderate pain (5 Point)</td>
<td>Severe pain (0 Point)</td>
</tr>
</tbody>
</table>

A total of 138 sets of vectors from the training set were input into the ten inputs of the neural network and trained. The training was performed 500 iterations. In each iteration, the labels were predicted and compared based on the 138 sets of vector data, and the weights were updated based on the comparison results.

**Softmax activation function**

In this study, BP neural network was used to consider the softmax function as the activation function, which can map the input features into a probability vector such as in the two classification tasks performed in this study. The softmax activation function mapped it into a two-dimensional probability vector, and then turned the larger probability value into 1 and the other into 0, and then output it. Softmax formula was as follows:

$$a_j = \frac{e^{z_j}}{\sum_{k=1}^{n} e^{z_k}}$$  \hspace{1cm} (1)

**Cross-entropy loss function**

The cross-entropy loss function, as the most commonly used loss function for neural networks, can accurately calculate the degree of difference between the theoretical and actual values of the model [34]. In general, the smaller the value of the loss function, the more the test results of the model reflect the actual values. It was also fed back to the neural network for correction. This loss function was used to check each parameter and compare the prediction vector with the known fact vector, which led to an improvement in the accuracy of the prediction [35].

**Use of Dropout to prevent overfitting**

Overfitting is a problem that can easily occur in neural network. The influence between different feature detectors can cause the neural network to overfit the data used for training but not for the newly added test data, which in turn makes the accuracy decrease. Dropout can ignore half of the feature detectors in each training batch (to let half of the hidden layer nodes have value 0) [36], i.e., let the activation value of a neuron during forward propagation with certain probability p to stop working, this approach may reduce the interactions between feature detectors (hidden layer nodes) and make the model more generalizable because it does not rely too much on certain local features. The occurrence of overfitting can be mitigated more effectively and regularization can be achieved by Dropout.

**Weight output**

After finishing the training, weight function was applied to output the weights of each input and each intermediate layer to obtain a vector of 10 groups and 32 items. After 500 iterations of training, the weights of each of the 10 data were calculated for the hidden layer, and the absolute values of all the hidden layer data were summed up to represent the relative weights of the data for the final output.

**Establishment of prediction model for surgical methods**

The trained neural network was used to construct the prediction model of the surgical method. The prediction of new data was performed by applying the predict function. In the predict function, a set of 10 parameters for a new patient was entered in the same order and the label prediction of the new data was performed using the trained neural network. The built model can then be applied for surgical method prediction (Figure 2).

**Batch processing and validation**

The function of reading in csv and outputting excel in Python were used to achieve batch processing. 45 sets of patient data from the test group were input into the csv file "input2.csv". These data was read in and input into the neural network using Python, batch prediction was performed, and the prediction results were output to a pre-set excel file, "prediction results.xlsx". At the same time, the actual surgical procedure used for those 45 patients was prepared and compared.

**Results**

**Weight analysis**

After training on the data set of 138 patients, the weights of the 10 parameters were output for each hidden layer neural node, and the absolute values were summed for each class of data to obtain the relative weight values Q1-Q10 for that parameter. Finally, Q1 to Q10 were normalized, and the normalization formula was as follow:

$$N_a = \frac{Q_a}{Q_1 + Q_2 + Q_3 + \ldots + Q_9 + Q_{10}} \times 100\%$$

The relative weights obtained for each parameter among the 10 parameters are shown in Table 2.
Construction of prediction model for surgical method

The neural network trained with 138 sets of data vectors was able to show the predicted surgical approach by predict function. At the same time, the neural network implemented a batch processing by code. Using this function, the present study verified the accuracy of the prediction model constructed by this neural network. The results were classified into four categories based on the surgical approach predicted by the model and the actual surgical approach: true TKA (actual TKA, predicted TKA), false TKA (actual UKA, predicted TKA), true UKA (actual UKA, predicted UKA), and false UKA (actual TKA, predicted UKA). Also, the percentage of data in each category was verified. This process had 50 cycles and each cycle will re-disrupt the training and testing groups and clear the neural network memory and retrain. For each loop, 45 data groups are used as test groups to verify the prediction results, and a total of 2250 predictions were made. After 50 loops, the number and percentage of each result were recorded as shown in Table 3.

The model prediction accuracy equaled the percentage of true TKA plus the percentage of true UKA, and the result was 92.97%.

Discussion

In this study, data related to KOA patients, including basic condition data, KSS score data, and X-ray imaging data were recorded, and range of motion of knee was measured based on wearable devices for 183 patients. A BP neural network was applied to analyze these parameters to derive their relative weights in the diagnosis. Also, the neural network was trained using these data to build a prediction model for surgical method selection.

When KOA patients are diagnosed in clinic, their basic conditions, KSS score indicators, and radiographic images are all different. Ultimately, the used surgical method varies as well. The surgical procedure chosen is influenced by multiple factors.

Whether the narrowing of the lateral compartment of the patient's knee on the X-ray image or not is extremely important for the final choice of surgical method. In this study, the prediction model showed that its percentage of 15.37% of the 10 items was significantly higher than the other parameters. Thus, based on the data, it can be determined that KOA patients with the narrowing of the lateral compartment were more inclined to accept the operation of TKA, while patients without the narrowing of the lateral compartment were more favored for UKA. Whether the narrowing of medial compartment or not, with or without sclerosis, although less influential than whether the narrowing of the lateral compartment or not in the choice of surgery, also has a weighting ratio of more than 10%, stronger than ROM, gender, age, etc.

The relative weight of gender was 9.39% of the 10 items, so the patient's gender influenced to some extent the surgical method that will be used. Physicians should be aware of male patients and take them into account as a factor in their diagnosis.

The age of the patients in this study varied widely, from 40 to 90 years old, and what can be determined from the age factor was that the older the patient, the more likely they were to choose TKA, which may be related to their older age and more severe joint wear.

The patient's weight, as it relates to the amount of supported force on the patient's knee joint, had an impact on the diagnosis of the disease and therefore the choice of surgical procedure. However, the results showed that it was not a decisive factor in the choice of surgical procedure.

BMI, as the patient's basic physical index, has some influence on the patient's surgical choice. The surgeon may pay attention to the patient's BMI during diagnosis as a reference factor for surgical selection.

ROM is a more objective data to determine the functional condition of the patient's knee joint, but its final influence on

Table 2: Relative weight of each parameter.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Gender</th>
<th>Age</th>
<th>Weight</th>
<th>BMI</th>
<th>ROM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative</td>
<td>0.0939</td>
<td>0.0989</td>
<td>0.0826</td>
<td>0.0936</td>
<td>0.0978</td>
</tr>
<tr>
<td>weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Pain score</td>
<td>Osteophytes</td>
<td>Sclerosis</td>
<td>medial space</td>
<td>lateral space</td>
</tr>
<tr>
<td>Relative</td>
<td>0.0834</td>
<td>0.085</td>
<td>0.0987</td>
<td>0.1125</td>
<td>0.1537</td>
</tr>
<tr>
<td>weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Test results of prediction model accuracy.

<table>
<thead>
<tr>
<th>Prediction types</th>
<th>True TKA</th>
<th>True UKA</th>
<th>False TKA</th>
<th>False UKA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>1203</td>
<td>889</td>
<td>158</td>
<td>0</td>
</tr>
<tr>
<td>Average percentage (%)</td>
<td>0.5346</td>
<td>0.3951</td>
<td>0.0702</td>
<td>0</td>
</tr>
</tbody>
</table>
the choice of surgical modality was 9.78% out of 10 items. It was almost consistent with gender, age, and BMI. And pain scores were weighted lower.

Bone osteophytes or osteoproliferation, had less impact on the choice of applying UKA or TKA.

In the prediction of surgical method by the neural network established in this study, there were still prediction errors, and it could be found that the prediction made by the model never occurred for false UKA, but the error always occurred for false TKA. i.e. patients who should accept UKA, the model predicted that they would accept TKA. The reason for this phenomenon was that in the training data, there were more patients with TKA, compared to those with UKA. The insufficient amount of data for UKA resulted in the trained model being more biased toward selecting TKA for surgery. Most patients with KOA were female. In the weighting analysis, the number of female patients in this study was also more than the number of male patients, and the majority of patients with KOA in the moderate and severe stages had bone osteophytes, therefore, it may lead to a lower weighting of gender and bone osteophytes for the surgical outcome compared to the actual situation.

The prediction model established in this study had the great application value and still had many aspects for improvement. Firstly, the quality and quantity of data may be improved, the relevance between patient data and labels can be strengthened, case data discussed by multiple experts can be used as training data, and the number of cases can be significantly expanded, and the generalization and accuracy of the model will be qualitatively improved. Secondly, the model in this study used only ten covariates for prediction, it is possible to increase the number of data as factors, such as other factors in X-ray images, knee varus angle in KSS scores, and even medical history, can be added to the model to improve the model accuracy. At the same time, functions such as image recognition can be added and interfaced with the App. The diagnostic methods and processes for doctors and patients will be substantially improved, and with mature technology, even automatic diagnosis can be achieved. Finally, this research model can be applied to other diseases by building similar prediction or diagnosis models based on patient case data of other diseases.

In future work, the number of cases will be increased and the quality of the data will be improved to optimize the neural network as a way to improve its usefulness.

Conclusion

The accuracy of the established surgical method prediction model was 92.97%. Physicians should use the narrowing of lateral compartment as a key indicator in the diagnosis. Also the narrowing of the medial compartment or not and the presence or absence of sclerosis should be used as an auxiliary judgment basis for the choice of surgical method. Bone osteophytes was difficult to use as a basis for judgment because of its small effect. ROM as a more important index, doctors must mainly refer to it. As for the degree of pain, it was difficult to use it as a decisive basis for judgment because different patients felt it differently. Age, gender, and BMI were all related to the final surgical method, so doctors should pay attention to these three values when diagnosing patients, while weight was not related to the patient's surgical method, and can only be considered as an appropriate reference.

Acknowledgement

We thank the subjects participating this study and also thank the funding of Beijing Natural Science Foundation of China (Grant 7222300).

Author Contributions

Zhang Haohua and Wang Jiaying contributed equally to this work and are first co-authors.

Conflict of Interest

None.

References


